

Edris M. Alkozai, Bakhtawar K. Mahmoodi, Johan Decruyenaere, Robert J. Porte, Annemieke Oude Lansink-Hartgring, Ton Lisman and Maarten W. Nijsten\*

## Systematic comparison of routine laboratory measurements with in-hospital mortality: ICU-Labome, a large cohort study of critically ill patients

<https://doi.org/10.1515/cclm-2016-1028>

Received November 8, 2016; accepted November 23, 2017; previously published online January 8, 2018

### Abstract

**Background:** In intensive care unit (ICU) patients, many laboratory measurements can be deranged when compared with the standard reference interval (RI). The assumption that larger derangements are associated with worse outcome may not always be correct. The ICU-Labome study systematically evaluated the univariate association of routine laboratory measurements with outcome.

**Methods:** We studied the 35 most frequent blood-based measurements in adults admitted  $\geq 6$  h to our ICU between 1992 and 2013. Measurements were from the first 14 ICU days and before ICU admission. Various metrics, including variability, were related with hospital survival. ICU-based RIs were derived from measurements obtained at ICU discharge in patients who were not readmitted to the ICU and survived for  $>1$  year.

**Results:** In 49,464 patients (cardiothoracic surgery 43%), we assessed  $>20 \cdot 10^6$  measurements. ICU readmissions, in-hospital and 1-year mortality were 13%, 14% and 19%, respectively. On ICU admission, lactate had the strongest relation with hospital mortality. Variability was independently

related with hospital mortality in 30 of 35 measurements, and 16 of 35 measurements displayed a U-shaped outcome-relation. Medians of 14 of 35 ICU-based ranges were outside the standard RI. Remarkably,  $\gamma$ -glutamyltransferase (GGT) had a paradoxical relation with hospital mortality in the second ICU week because more abnormal GGT-levels were observed in hospital survivors.

**Conclusions:** ICU-based RIs for may be more useful than standard RIs in identifying ICU patients at risk. The association of variability with outcome for most of the measurements suggests this is a consequence and not a cause of a worse ICU outcome. Late elevation of GGT may confer protection to ICU patients.

**Keywords:**  $\gamma$ -glutamyltransferase; critical care; lactate; outcome; reference interval; variability.

## Introduction

Critically ill patients often have laboratory measurements that are abnormal when compared with standard reference intervals (RIs). The use of standard RIs derived from healthy persons to assess disease severity or identify complications may sometimes be inappropriate in patients in the intensive care unit (ICU). In order to identify specific pathophysiological mechanisms or to develop multivariate predictive models for critically ill patients, many ICU studies have evaluated the relation between selected measurements and outcome [1–5]. Depending on which measurements are selected or which clinical phase is considered of interest, measurements were from patients who may be outside the standard RI.

It has been proposed to define RIs for specific patient groups such as hospitalized patients [6]. Moreover, the implicit assumption that a more deranged measurement will be associated with a worse clinical situation may not always be correct.

The goals of the ICU-Labome study were to comprehensively evaluate the univariate relation of regular laboratory measurements with outcome and to identify “ICU-based” RIs, derived from ICU patients who were discharged and had no major complications as reflected

\*Corresponding author: Maarten W. Nijsten, Department of Critical Care, University of Groningen, University Medical Center Groningen, Hanzeplein 1, 9700 RB Groningen, The Netherlands, Phone: 00-31-50-3616161, Fax: 0031-50-3615644, E-mail: m.w.n.nijsten@umcg.nl

Edris M. Alkozai: Department of Critical Care, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; and Department of Surgery, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Bakhtawar K. Mahmoodi: Department of Cardiology, Sint Antonius Hospital, Nieuwegein, The Netherlands

Johan Decruyenaere: Department of Intensive Care Medicine, Ghent University Hospital, Ghent, Belgium

Robert J. Porte and Ton Lisman: Department of Surgery, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Annemieke Oude Lansink-Hartgring: Department of Critical Care, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

by no ICU readmission and 1-year survival. Parameter-specific *a priori* assumptions or multivariate models such as APACHE [2, 3] were not the scope of this study. We also checked for potential U-shaped relations with outcome and whether measurement variability was associated with outcome, as observed for glucose and potassium [7–9].

## Materials and methods

### Supplementary material

Given the scope of this study, some methods and the majority of the (intermediate) results of analyses are reported in an extensive Supplementary Material File (SMF.pdf), so we could consistently report the various aspects of all the 35 measurements.

### Patients and outcome

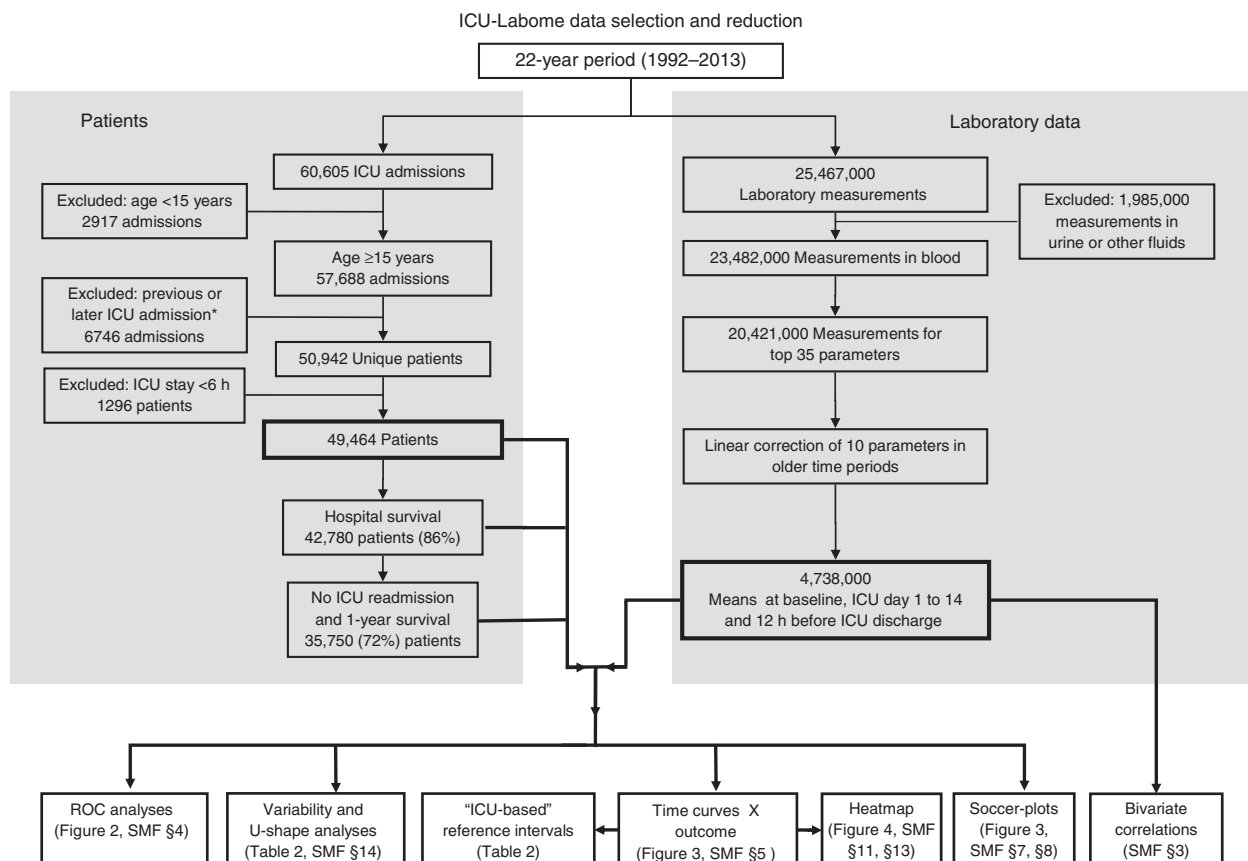
From 1992 through 2013, all laboratory measurements from all patients admitted to our tertiary 44-bed ICU in a University Hospital

were evaluated (Figure 1). Patients  $\geq 15$  years were included and data were directly anonymized before further analysis. The type of ICU admission was recorded, and when patients were admitted multiple times to the ICU, the first ICU admission of the last hospital admission was used. Patients who stayed  $< 6$  h at the ICU were excluded.

In-hospital mortality was the main outcome measure. ICU mortality, ICU readmission, lengths of stay and 1-year survival were also recorded. The hospital information system is periodically updated with national survival status information for patients treated in the ICU.

### Selection, correction and primary reduction of measurements

We selected the 35 most frequently assessed laboratory measurements in the blood during the first 14 days of ICU stay, or directly preceding ICU admission. The numbers 35 and 14 were arbitrarily chosen to cover the vast majority of measurements. Derived values such as base excess were excluded. The standard RIs used were those used by our central diagnostic laboratory in December 2013 (Table 1). We excluded only obviously impossible values, typically resulting from data entry or storage mistakes. This concerned negative, zero or extreme values for assays that could not possibly report such values. Some standard RIs were modified over the 22-year study period. We



**Figure 1:** Data flow.

Flowchart depicting selection of patients, selection of the top 35 laboratory parameters and subsequent data reduction and analysis.

\*In case of multiple ICU stays, only the first ICU stay of the last hospital admission was used.

**Table 1:** Laboratory measurements evaluated.

Abbreviation	Measurement	Unit	sRI <sup>1</sup>	Type of relation with mortality <sup>2</sup>	Variability and mortality <sup>4</sup>	'ICU-based' RI: median (IQR) <sup>5</sup>	Median of ICU-based RI relative to sRI <sup>1</sup>
ALAT	Alanine aminotransferase	U/L	<45	Monotonic	+++	30 (20–48)	in sRI <sup>5</sup>
Alb	Albumin	g/L	35–50	U-shaped	+	28 (24–32)	<sRI <sup>5</sup>
Amy	Amylase	U/L	<107	Monotonic	+	67 (38–115)	in sRI
AP	Alkaline phosphatase	U/L	<115	Monotonic	+	55 (43–75)	in sRI
aPCO <sub>2</sub>	Arterial pCO <sub>2</sub>	kPa	4.6–6.0	U-shaped	++	5.0 (4.7–5.5)	in sRI
apH	Arterial pH		7.35–7.45	U-shaped	++	7.40 (7.37–7.43)	in sRI
aPO <sub>2</sub>	Arterial PO <sub>2</sub>	kPa	9.5–13.5	U-shaped	++	12.4 (10.6–15.0)	in sRI
APTT	Activated partial thromboplastin time	s	23–33	Monotonic	+	27 (25–31)	in sRI
ASAT	Aspartate aminotransferase	U/L	<35	Monotonic	+	40 (28–62)	>sRI <sup>5</sup>
aSatO <sub>2</sub>	Arterial oxygen saturation		0.96–0.99	Monotonic	+	0.98 (0.96–0.99)	in sRI
Bic	Bicarbonate	mmol/L	21–25	Monotonic	+	23 (21–25)	in sRI
Ca	Calcium (total)	mmol/L	2.20–2.60	U-shaped	++	1.96 (1.84–2.08)	<sRI
CK-MB	Creatine MB kinase	U/L	<5	Monotonic	+	14 (8–24)	>sRI
CKT	Total creatine kinase	U/L	<170	U-shaped <sup>3</sup>	–	215 (121–415)	>sRI
Cl	Chloride	mmol/L	97–107	U-shaped	+++	107 (104–110)	>sRI
Creat	Creatinine	μmol/L	50–110	Monotonic	++	68 (57–83)	in sRI
CRP	C-reactive protein	mg/L	<5	Monotonic	–	57 (23–107)	>sRI
DBI	Direct bilirubin	μmol/L	<5	Monotonic	+	4 (2–7)	in sRI
GGT	γ-Glutamyltransferase	U/L	<55	Monotonic	–	36 (18–86)	in sRI
Glu	Glucose	mmol/L	4.0–6.0	U-shaped	++	7.4 (6.3–8.9)	>sRI
Hb	Hemoglobin	mmol/L	7.7–10.6	U-shaped	+++	6.4 (5.8–7.2)	<sRI
Ht	Hematocrit	mmol/L	0.36–0.52	U-shaped	++	0.30 (0.28–0.3)	<sRI
K	Potassium	mmol/L	3.5–5.0	U-shaped	++	4.3 (4.0–4.6)	in sRI
Lac	Lactate	mmol/L	0.5–1.5	Monotonic	–	1.0 (0.8–1.4)	in sRI
LDH	Lactate dehydrogenase	U/L	<248	Monotonic	++	297 (213–408)	>sRI
Leuko	Leukocyte count	10 <sup>9</sup> /L	4–10	U-shaped <sup>3</sup>	++	12.2 (9.7–15.3)	>sRI
Mg	Magnesium	mmol/L	0.70–1.00	Monotonic	–	0.95 (0.82–1.11)	in sRI
Na	Sodium	mmol/L	135–145	U-shaped	++	137 (135–139)	in sRI
P	Phosphate	mmol/L	0.70–1.50	U-shaped <sup>3</sup>	++	1.02 (0.85–1.20)	in sRI
PLC	Platelet count	10 <sup>9</sup> /L	150–350	U-shaped	++	180 (138–239)	in sRI
PT	Prothrombin time	s	9.0–12.0	Monotonic	+	11.6 (10.8–12.7)	in sRI
TBI	Total bilirubin	μmol/L	<17	Monotonic	+	10 (7–16)	in sRI
TP	Total protein	g/L	60–80	U-shaped	+	53 (47–59)	<sRI
Trop	Troponin	ng/L	<14	Monotonic	+	51 (15–124)	>sRI
Urea	Urea	mmol/L	2.2–7.5	Monotonic	+	6.4 (4.9–8.7)	in sRI

Key characteristics and results of 35 laboratory parameters evaluated in the ICU-Labome study. <sup>1</sup>sRI: standard reference interval as provided by the central laboratory. <sup>2</sup>All parameters obtained showed a univariate relation with in-hospital mortality, and for 13 parameters this relation was U-shaped when a quadratic function was used. <sup>3</sup>For three additional parameters (CKT, leuko and P), a U-shape was only found with a more complex polynomial function (SMF §14). <sup>4</sup>The relation of standard deviation (SD) as a measure of variability with in-hospital mortality was classified as detailed in the methods section from '–' through '+++'. Thus, an independent association of variability with in-hospital mortality was present for the majority of parameters. Furthermore, in 16 of the 35 parameters, SD had a stronger relation with outcome than the mean (i.e. '++' and '+++'). <sup>5</sup>Medians with interquartile ranges (IQR) obtained at discharge in patients who achieved good outcome (i.e. no readmission to the ICU and survival >1 year) were be used as 'ICU-based' RIs. <sup>5</sup>"in sRI", "<sRI" and ">sRI", respectively, denote that the median ICU-based value provided in the previous column is within, below or above the standard reference interval (sRI). For 14 of the 35 measurements, the median of the "ICU-based IQR" was found to be outside the standard reference interval.

verified for measurements whether abrupt time-dependent long-term changes had occurred, and if required we performed a linear correction to adjust such measurements (SMF §9).

ICU days 1 through 14 were determined in 24-h blocks counting from the date and time of ICU admission. When patients had

multiple measurements within the same ICU day, the mean was calculated before further analysis, except for the analysis of variability. For measurements directly preceding ICU admission, the mean over the 120 h before ICU admission was considered the baseline value.

## Bivariate correlations

We performed bivariate linear regression analysis for the 35  $\cdot (35 - 1)/2$  measurement pairs to identify strongly related or obviously redundant measurements. Data from all patients and ICU days were pooled for this analysis.

## Area under the receiver operating characteristics curve (AUROC)

To assess the univariate monotonic relation with outcome of the 35 measurements on ICU day 1, we determined the AUROC with in-hospital mortality as the dependent variable. Likewise, AUROCs were calculated for all measurements for ICU day 2, the change from ICU days 1 to 2, and for the 12-h window before ICU discharge. Note that measurements with a known U-shaped relation (see below) with outcome will do relatively poorly in AUROC analysis because this analysis assumes a monotonic relation of a test value with outcome.

## Variability (SD) and outcome

The relation of variability for the 35 measurements obtained up to ICU day 14 with in-hospital mortality was assessed. Similar to how variability has been determined for glucose [8] or potassium [9], variability of each measurement for each patient was defined as the standard deviation (SD) of all measurements, including multiple measurements on the same day, obtained during the ICU stay. Whether the SD was relevantly associated with outcome was assessed by performing logistic regression analysis with in-hospital mortality as dependent and the measurement's mean and SD as independent factors.

SD's association with outcome was classified as follows:

- ‘-’ No association;
- ‘+’ Both SD and mean were associated with outcome, but mean had a stronger association;
- ‘++’ Both SD and mean were associated with outcome, but SD had a stronger association;
- ‘+++’ Only SD was associated with outcome.

## U-shaped relation with outcome

For some measurements (e.g. sodium), both low and high levels are associated poor outcome. To assess the presence of such a U-shaped relation with outcome, logistic regression analysis was used with the individual mean measurement ( $\bar{x}$ ) and the squared mean measurement ( $\bar{x}^2$ ) on ICU day 1 as independent parameters and in-hospital mortality as dependent parameter. A U-shaped relation was considered present when both coefficients of the quadratic function were significant and the parabola had a minimum (i.e. lowest point of the U-curve) that was situated within the 10%–90% range of all individual means. A more complex polynomial fractional function with the terms  $\bar{x}^{-2}$ ,  $\bar{x}^{-1}$ ,  $\bar{x}^{0.5}$ ,  $\bar{x}$  and  $\bar{x}^2$  was also explored in a similar manner.

## Time course of medians and soccer plots

We compared the time course of the medians  $\pm$  interquartile range (IQR) of the 35 laboratory measurements between hospital survivors and hospital non-survivors.

‘Soccer’ plots [10] were constructed to provide additional graphical information on the distribution of values in relation with the standard RI, as a function of ICU day for in-hospital survivors and non-survivors. Values within the standard RI are green, whereas yellow, orange and red reflect values both below and above the standard RI, according to detailed criteria (SMF §6).

The observed distribution of laboratory measurements may have structurally changed over the years, either because changes in the ICU (e.g. glucose control) or because of changes in the laboratory (e.g. a modified assay for albumin). For visual recognition of structural changes in the distribution of the 35 reported measurements, soccer plots were also made for the 1992–2002 and 2003–2013 periods.

## ‘ICU-based’ RIs

We used ICU patients who were not readmitted to the ICU and survived for  $\geq 1$  year to generate ‘ICU-based’ reference ranges [6]. The clinical period that was used to obtain these laboratory measurements was the last 12 h of the ICU admission before ICU discharge. To obtain conservative estimates, IQRs (i.e. P25%–P75%) were determined, and not the P2.5% and P97.5% as is usually done for standard RIs derived from a population of normal individuals.

## Heat map

One would expect that when laboratory measurements are deranged, i.e. below or above the standard RI, that in the sickest patients such measurements would be furthest from the standard RI. To verify this, we examined when median levels in patients who died during the same hospital stay were more deranged than in patients who survived. A heat map was constructed to summarize in a single figure for all 35 median measurements at all time points whether they were more deranged (orange or red), similarly deranged (gray) or less deranged (green or dark green) in non-survivors than survivors.

## Statistics

Distributions were compared with the  $\chi^2$ -test and medians with the Mann-Whitney U-test. Logistic regression analyses were used only for one of the 35 measurements at the time. As construction of multivariable predictive models was not a goal of this study, no models using multiple lab measurements were used. In order to assess the contribution the variability of a parameter  $\bar{x}$  (i.e. SD) to outcome, the mean and SD of  $\bar{x}$  were entered as factors into logistic regression analysis. Similarly to assess a U-shaped relation,  $\bar{x}$  and  $\bar{x}^2$  were entered (SMF §14). To identify more complex U-shaped relations, we also performed the same analyses with the terms  $\bar{x}^{-2}$ ,  $\bar{x}^{-1}$ ,  $\bar{x}^{0.5}$ ,  $\bar{x}$  and  $\bar{x}^2$  (fractional polynomials [11]) with conditional backward elimination of terms with a  $p \leq 0.10$  for entry and  $p \geq 0.05$  for removal of

terms. Plots of all 35 quadratic (SMF §14.1a to §14.35a) and fractional polynomial fits (SMF §14.1b to §14.35b) were generated to enhance interpretation. SPSS version 23 from IBM was used for all analyses and Microsoft Excel and Powerpoint 2010 were used for graphical representations.

### Ethics approval and consent to participate

This study was approved by our institutional review board (IRB), the “Medisch ethische toetsingscommissie of the University Medical Center Groningen” (METC 2014.264). Because only anonymized data that had been obtained during routine care were analyzed and no additional sampling or interventions were performed for this retrospective study, the IRB considered informed patient consent not necessary.

## Results

The overall major data reduction and synthesis stages with the corresponding numbers of patients and measurements are shown in Figure 1.

### Patients and outcome

Over the study period, there were 60,605 ICU admissions (Figure 1). Excluded were 18% (<15 years: 5%; ICU re-admission: 11%; ICU stay <6 h: 2%). The mean  $\pm$  SD age of the selected population was  $60 \pm 16$  years, 37% were females and 18% were admitted from the emergency department. The largest admission category was cardiothoracic surgery (Table 2).

Mean ICU and hospital length of stay were  $4.3 \pm 9.6$  and  $18 \pm 21$  days, and the frequency distributions of the ICU length of stay did not markedly change from 1992–2002 to 2003–2013 (SMF §2). Thirteen percent of the patients were readmitted to the ICU during the same hospital stay. ICU, in-hospital and 1-year mortality were 11%, 14% and 19%, respectively. Non-survivors had longer ICU stays than survivors (SMF §2.2). Of all studied ICU patients, 72% were not readmitted and were still alive after 1 year.

### Selection, correction and primary reduction of measurements

Of  $23 \cdot 10^6$  blood measurements recorded,  $20 \cdot 10^6$  (87%) were included (Figure 1). The frequency distribution of the number of mean measurements per ICU day is provided

**Table 2:** Patient characteristics.

No. of patients	49,464
Age, years (SD)	60 (16)
Sex, male	63%
Admission from emergency department	18%
Type of admission	
Cardiothoracic surgery	42.5%
Abdominal, vascular and miscellaneous surgery	14.5%
Neurosurgery	11.2%
Medical	5.8%
Trauma	3.8%
Transplantation	1.8%
Miscellaneous	20.4%
Mean (SD) ICU stay <sup>a</sup> , days	4.3 (9.6)
Median (IQR) ICU stay <sup>a</sup> , days	1.0 (0.8–3.1)
Mean (SD) hospital stay, days	18 (21)
Median (IQR) hospital stay, days	11.6 (7.2–21.0)
ICU readmission	13%
ICU mortality	11%
In-hospital mortality	14%
1-year mortality	19%
Patients in reference group <sup>b</sup>	72%

ICU, intensive care unit; IQR, interquartile range; SD, standard deviation. <sup>a</sup>ICU stay includes ICU stay for readmissions. <sup>b</sup>Patients in this group had no ICU readmission and were alive after 1 year. This group was used to calculate ICU based reference intervals.

in SMF §1.1. For 10 measurements, linear adjustments were made (SMF §9.2) due to evident structural changes in reported laboratory values over the years. In several cases (e.g. albumin), the running means of the median laboratory values still reflected considerable gradual changes even after these corrections (SMF §10).

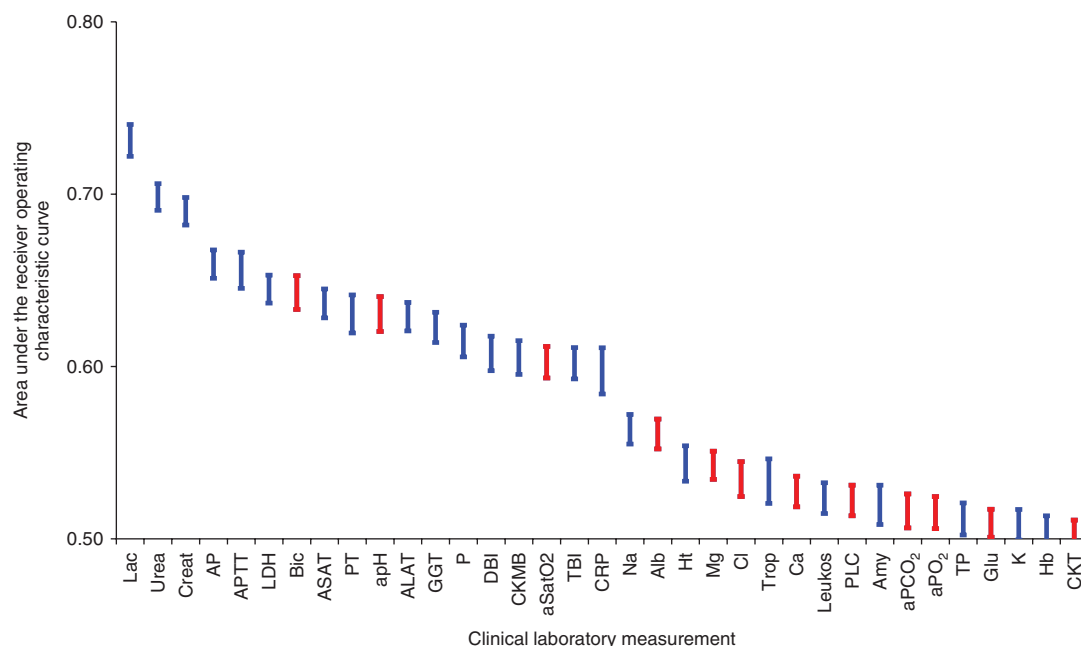
### Bivariate correlations

Many of the 595 measurement pairs examined showed a positive  $R$  (SMF §3). The highest  $R$ s were +0.96 and +0.97 for the hemoglobin, hematocrit and direct bilirubin, total bilirubin pairs, respectively. Fewer negative correlations were observed and these were not as marked as positive correlations. The most negative  $R$  was  $-0.50$  for the  $\text{apH}$ , lactate pair.

### AUROC

Thirty-two out of 35 laboratory measurements had an AUROC  $>0.50$ . Lactate on ICU day 1 showed the strongest predictive power for in-hospital mortality (AUROC of 0.731; 95% CI: 0.722–0.740; Figure 2) as well as lactate on





**Figure 2:** AUROC analysis on ICU day 1.

Univariate relation of laboratory parameters on ICU day 1 with in-hospital mortality as assessed by area under the receiver operating curve (AUROC). We assessed the univariate prognostic relevance of 35 studied laboratory variables on ICU day 1. All parameters had an AUC higher than 0.5, except for the total creatine kinase (CKT), hemoglobin (Hb), and potassium (K). Lactate (Lac), urea, and creatinine were the strongest predictors of in-hospital mortality at ICU day 1. Blue color reflects a positive relation with mortality; red color reflects an inverse relation with mortality. Note that the AUROC will do poorly for variables with a U-shaped outcome relation. AUROC are displayed with 95% confidence intervals.

ICU day 2. AUROC data for ICU day 1 and 2, the change from ICU days 1 to 2, and at <12 h before ICU discharge are provided in SMF §4.

## Variability (SD) and outcome

For 30 of the 35 measurements, SD had an independent relation with in-hospital mortality (Table 1). Moreover, in 16 of these measurements, the SD had a stronger relation with outcome than the mean (alanine aminotransferase [ALAT], arterial partial CO<sub>2</sub> pressure [aPCO<sub>2</sub>], arterial pH [aph], aPO<sub>2</sub>, calcium, chloride, creatinine, glucose, hemoglobin, hematocrit, potassium, lactate dehydrogenase [LDH], leukocyte count, sodium, phosphate and platelet count; Table 1).

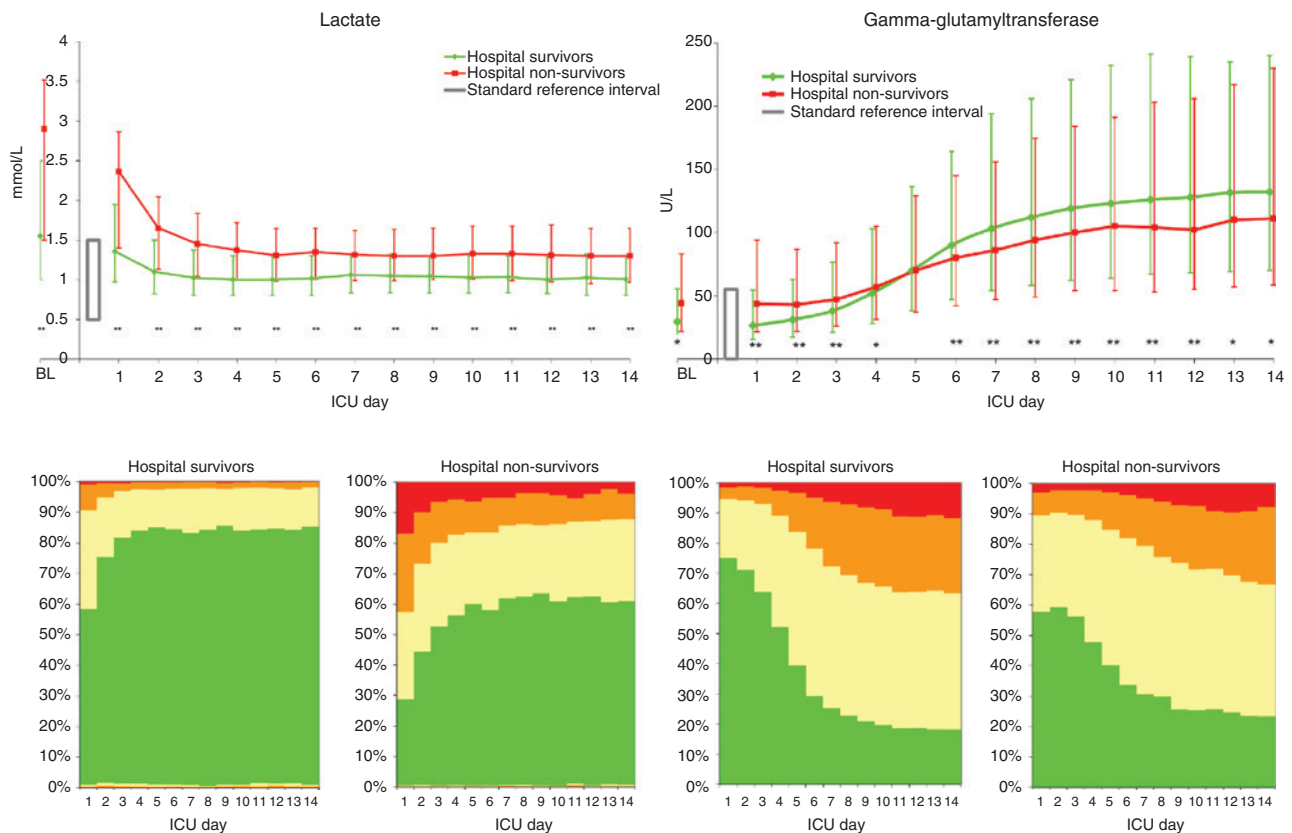
## U-shaped relation with outcome

At ICU day 1, all 35 parameters evaluated showed a univariate relation with in-hospital mortality, and in 16 of them (Table 1), this relation was U-shaped.

## Time course of medians and soccer plots

Time courses of the medians of all 35 measurements for hospital survivors and non-survivors and the associated p-values are provided in SMF §5.1–§5.35 and §11, respectively. Two of the most remarkable time courses, lactate and  $\gamma$ -glutamyltransferase (GGT), are shown in Figure 3. Lactate was considerably higher at all time points in non-survivors ( $p < 10^{-21}$  to  $p < 10^{-300}$ ) and GGT was initially higher ( $p < 3 \cdot 10^{-156}$ ) but subsequently lower in non-survivors ( $p < 5 \cdot 10^{-7}$ ).

The soccer plots (SMF §7) demonstrate that for some parameters, most measurements were outside the standard RI (e.g. albumin, SMF §7.2), whereas for other parameters (e.g. potassium, SMF §7.23), most were inside the standard RI but extreme values predominantly occurred in non-survivors. Comparison of soccer plots between 1992–2002 and 2003–2013 shows that many changes over the first 14 ICU days have remained similar over the past decades (SMF §8). However, some exceptions are evident. Albumin, aPCO<sub>2</sub>, chloride, hemoglobin, hematocrit and sodium [10] became more deranged during ICU stay in the 2003–2013 period, whereas glucose became less deranged. These effects might be explained by well-known changes



**Figure 3:** Time courses of lactate and  $\gamma$ -glutamyltransferase (GGT) and outcome.

Changes in lactate and GGT during the first 14 ICU days graphically represented in two different manners. The two upper panels show medians with interquartile ranges for survivors (green) and non-survivors (red). Lactate showed the strongest discrimination between survivors and non-survivors for all 35 measurements that were evaluated. At all time points non-survivors had higher lactate levels. GGT was initially significantly lower for in-hospital survivors, but it increased to levels that are higher than and more out of range than those of non-survivors from ICU day 6 onwards. Gray bars denote standard reference intervals; \*\* $p < 0.0005$ ; \* $p < 0.01$ . The lower panels are soccer plots with green areas denoting measurements in the standard reference ranges, and yellow, orange and red denoting progressively more deranged values (see detailed description SMF §6). The soccer plot also indicates after ICU day 5 that survivors had more deranged GGTs than non-survivors.

in therapeutic strategies concerning which laboratory deviations are deemed acceptable.

A clear feature of some parameters is their complex time dependency over the first 14 ICU days. Most extreme was the quadriphasic behavior of the leukocyte count (SMF §5.26, §7.26, §8.26). Although the leukocyte count was consistently higher in non-survivors, it can be appreciated that these time fluctuations diminish leukocytosis' association with prognosis [12].

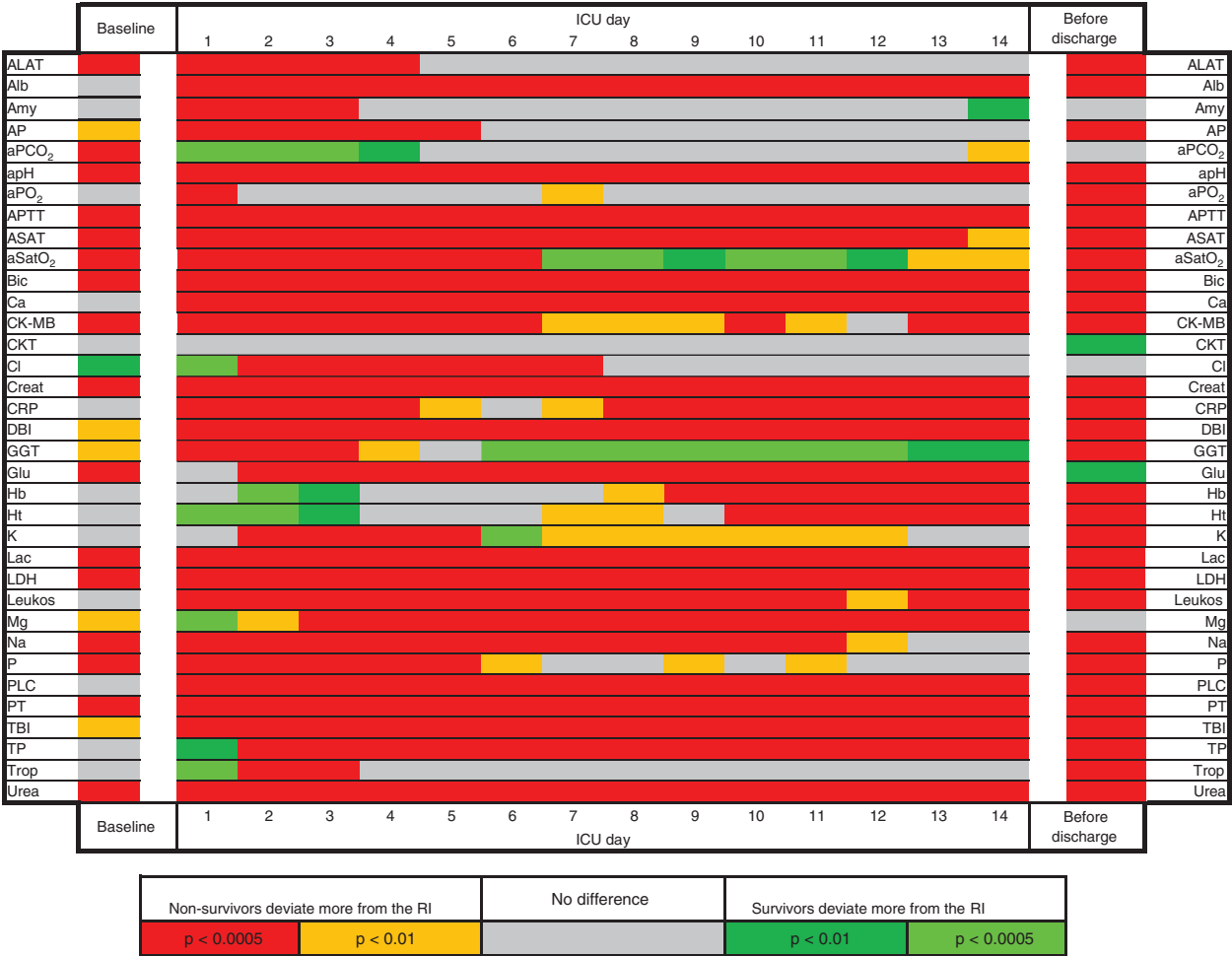
## 'ICU-based' RIs

When the median IQR laboratory values observed within 12 h of ICU discharge in patients who were not readmitted to the ICU and survived 1 year were considered as ICU-based IQRs, for 14 of the 35 laboratory measurements the median value of these ICU-based IQRs was outside the

standard RI (albumin, aspartate aminotransferase [ASAT], calcium, CK-MB, total CK, chloride, CRP, glucose, hemoglobin, hematocrit, LDH, leukocyte count, total protein and troponin; Table 1). Thus, many "abnormal" laboratory measurements before the discharge from the ICU were not "abnormal" from a prognostic point of view.

## Heat map

The heat map (Figure 4) indicates that many parameters are more deranged (i.e. red) at most or all time points (albumin, apH, APTT, bicarbonate, calcium, creatinine, lactate, PT and urea) in non-survivors with underlying  $p$ -values as low as  $<10^{-300}$  (SMF §11). Several parameters did not show differences at most time points, and few showed a 'paradoxical' difference (light and dark green) with a more deranged value in survivors. Between



**Figure 4:** Heat map. For the 35 parameters studied, red and orange in this heat map indicate that derangements (i.e. distance from official reference interval) for in-hospital non-survivors were larger than derangements for survivors (red:  $p < 0.0005$ ; orange:  $p < 0.01$ ). Gray indicates no significant difference between non-survivors and survivors. Green indicates a ‘paradoxically’ larger derangement for survivors than non-survivors (light green:  $p < 0.0005$ ; dark green:  $p < 0.01$ ), as observed for a few early measurements (e.g. aPCO<sub>2</sub>) that show an early paradoxical relation with outcome. After prolonged ICU stay, only GGT shows a consistently more deranged level in survivors. In SMF §13.1 and §13.2 heat maps of cardiothoracic surgery patients and non-cardiothoracic surgery patients are shown.

ICU days 5 and 14, only GGT showed highly significantly more deranged values in survivors. This paradoxical relation was also present when only patients who stayed  $\geq 14$  days in the ICU were analyzed 9 (SMF §12.1). Likewise, it was also observed in subgroup analyses for the cardiothoracic surgery (43%) and non-cardiothoracic surgery (57%) groups (SMF §13), as well for the miscellaneous surgery, trauma, and medical categories (SMF §12.2).

Discussion

In this comprehensive analysis of the blood-based measurements, we found that for 14 of the 35 most used

parameters, the ICU-based RI that we determined was outside the standard RI. Furthermore, our results underscore that early lactate has the strongest predictive value for in-hospital mortality of 35 laboratory measurements that were assessed. Also, we found that for most of the 35 measurements, variability had an independent relation with outcome. Close examination of the time courses in survivors and non-survivors over the first 14 ICU days uncovered a unique characteristic of GGT compared to other laboratory measurements. In the second week of ICU admission, the GGT levels were more deranged in survivors compared to non-survivors.

This analysis of regularly performed laboratory measurements had no *a priori* assumptions about specific mechanisms. We only hypothesized that the



standard RIs may not be a reliable tool in assessing critically ill patients because many patients still have derangements upon the discharge from the ICU. In fact, it has previously been shown that adjusted RIs for the ICU population decrease false-positive results and increase true-negative results [13]. Many laboratory derangements do not endanger the ICU patient, do not necessitate further intervention and should not preclude discharge from the ICU. Other derangements – although not life-threatening – might require attention in short-term clinical management. Obviously, when an observed value falls within the IQR as found in discharged ICU patients with a relatively good outcome, this does not automatically imply that this value is associated with minimal risk. In order to assess the mortality risk associated with a specific laboratory parameter value, even more advanced analyses are required [14, 15].

Concerning lactate, it has become evident that stress and not hypoxia is the most important driver of hyperlactatemia, which may explain its unique association with outcome [16–18]. Recently, the Sepsis-3 consensus incorporated lactate into the clinical definition of septic shock [19]. However, in common scoring systems used to predict mortality in ICU patients [2–5], many laboratory variables of lesser prognostic power are incorporated, but lactate is not yet included. We believe our results further support the routine use and inclusion of lactate into future scoring systems. The ICU-based RI for lactate closely corresponds with the standard RI (Table 1); thus, optimal lactate levels in ICU patients approach reference levels in healthy individuals.

The bivariate correlations showed that the (hemoglobin, hematocrit) and the (total bilirubin, direct bilirubin) pairs were strongly correlated, confirming that hematocrit in most cases is redundant on top of hemoglobin [20, 21]. Likewise, the apparent unconditional ordering of both total bilirubin and direct bilirubin measurements will not provide much additional information compared to total bilirubin alone, at least for our overall patient group.

With regard to measurement variability, it is remarkable that this metric was related with outcome for 30 out of 35 measurements. Although there has been a considerable focus on glucose variability as a therapeutic goal, it should be noted that the variability of – among others – several blood gas parameters, sodium and potassium had a stronger relation with outcome than the mean of these parameters (Table 1). In our judgment, this suggests that higher variabilities in patients who do worse are a fundamental reflection of the clinical instability of such patients [7–9, 22]. Thus, unless a specific explanatory mechanism

indicates otherwise, higher laboratory parameter instability should be considered a consequence and not a cause of a worse clinical situation.

Regarding the U-shaped relation of 16 measurements with outcome (Table 1), there were few unexpected findings. All these measurements, with the possible exception of albumin, are known to manifest both pathological ‘hypo’ and ‘hyper’ states.

In our view, the most surprising finding of our study was the paradoxical relation of GGT with outcome. GGT is a key enzyme in modulating redox-sensitive (extra) cellular defenses against toxins [23–26]. It is constitutively expressed in several organs and it breaks down extracellular glutathione (GSH), which generates cysteine for intracellular *de novo* synthesis of GSH. Higher serum GGT plausibly reflects increased cellular GGT activity and serum GGT increases with chronic exposure to toxic metabolites. Apart from its known association with the use of several drugs and ethanol [23], chronically elevated GGT in apparently healthy persons has emerged as strong risk factor for cardiovascular disease [27]. Likewise, in patients with liver disease, elevated GGT is considered a marker of cholestasis together with indicators such as bilirubin and alkaline phosphatase. We observed earlier that secondarily elevated GGT was associated with increased survival rates after liver transplantation, liver resection and after surgical repair of a ruptured abdominal aortic aneurysm [28–30]. The current observation in a large cohort of critically ill patients as well as in patients with prolonged ICU stay (SMF §12.1) or in subgroups without known primary liver disease (e.g. cardiothoracic or trauma patients; SMF §12.2) supports the notion that transiently elevated GGT is crucially involved in protective mechanisms leading to better outcome of ICU. In patients with acetaminophen intoxication, restoration of GSH levels with acetylcysteine treatment is critical [31, 32] as acetylcysteine provides cysteine for intracellular *de novo* synthesis of GSH. The benefit of acetylcysteine for related conditions is still unproven [33]. The possibility that interventions which increase GGT levels, such as GGT-inducing drugs [23], might also confer protection has also not been tested.

We believe that this paradoxical GGT relation could be confirmed in many existing ICU databases but has probably been overlooked because most studies focus on the first few ICU days. At the very least, caregivers should realize that the development of an elevated GGT in the second ICU week is not worrisome and should not automatically lead to investigations into its cause.

With regard to the other measurements that were green on the heat maps, these differences all occurred early during ICU stay with only small absolute differences

(e.g. hemoglobin, chloride or total protein; SMF §5, §11). Late “paradoxical” effects were seen for only few parameters, such as total creatine kinase activity (CKT) at discharge (Figure 4). It has been demonstrated earlier that lower enzyme activity levels can reflect a worse outcome [34]. However, comparison of heat maps of cardiothoracic surgery patients with other patients (SMF §13.1 and §13.2) indicates that this is a confounding effect from the large number of elective cardiothoracic surgery patients that have elevated postoperative CKT activity but have a lower mortality compared to other patients. It should also be remarked that the heat maps have many gray zones that reflect phases when abnormalities may be present but that do not discriminate for outcome.

A large body of ICU literature addresses the association of deviating laboratory measurements and outcome [1–5, 14, 15]. Our observations fully corroborate this because patients who did not survive the hospital admission had far more deranged laboratory measurements (Figure 4). However, our results may also have some practical consequences as they clearly show that the association of laboratory derangements with in-hospital mortality may sometimes be time dependent. Also, ICU care providers should realize that an abnormally high GGT may not necessarily be a cause for alarm or specific diagnostic procedures. In addition, our data underscore that many laboratory derangements in critically ill patients should not be considered “abnormal” from the prognostic point of view since derangements of 14 out of 35 laboratory measurements at ICU discharge were not associated with poor outcome (i.e. readmission to the ICU or deceased within 1 year).

Our study has a several limitations, some resulting from its retrospective design. We did not use commonly used ICU severity scores such as APACHE [2, 5] because these were not fully available. Many scoring systems such as APACHE use the most extreme daily value, not the daily mean as we did. Although using extremes often will improve a model’s predictive value, this involves additional arbitrary choices. To avoid this, we opted for the daily mean for all 35 measurements. We only performed univariate analysis of laboratory measurements with outcome as it was not our aim to create multivariable predictive models. Multivariable models can easily suffer from incorrect fits, including overfitting [35]. The way we defined the patient reference group (no ICU readmission and 1-year survival) may be considered arbitrary. However, we wanted to use a straightforward and reproducible criterion to identify the patients with a relatively good outcome [6]. Including 1-year survival served to exclude patients who did poorly post ICU discharge but were not readmitted to the ICU for various reasons. Covering more than two decades, many potential confounding

changes in measurement or therapy may have occurred, although data were at least partially adjusted for time effects (SMF §8, §10). The 1992–2002 vs. 2003–2013 soccer plots underscore that apart from the exceptions noted earlier, the behavior of most parameters has not essentially changed. During the ICU stay differences in the frequency of ordering measurements between certain patient groups may also have led to some bias, although no large differences in sampling between survivors and non-survivors were observed (SMF §1.2, §2.2). We do believe the extensive study period increased the robustness and external validity of many of our observations.

Despite our goal to analyze without *a priori* hypotheses, a number of choices had to be made, such as the limit of 35 measurements and the selected time windows. We chose to depict time courses of survivors and non-survivors over the 14 ICU days for all ICU patients, and not only those who had an ICU stay of  $\geq 14$  days. A majority of the patients were surgical patients with cardiothoracic surgery as a large subgroup (43%). However, the GGT-effect was also observed in other patient groups (SMF §12) and a heat maps of cardiac and non-cardiac surgery patients showed many similarities (SMF §13.1 and §13.2). Likewise, the time-dependent behavior of GGT was comparable between all patients and patients with an ICU stay  $\geq 14$  days (SMF §5.19 and §12.1).

We believe that our results represent clinically relevant information for the assessment of critically ill patients. Thus, we are confident that the observations can also be reproduced in other ICU cohorts. ICUs possessing large cohorts stored in integrated patient database management systems should also be able to explore relations between treatment and subsequent laboratory changes, such as an elevated GGT.

## Conclusions

The use of ICU-based RI’s instead of standard RI’s may decrease the level of uncertainty in clinical decision making. The widespread association of parameter variability with outcome makes it doubtful whether reducing variability of specific parameters is a useful therapeutic goal. Moreover, late elevation in GGT apparently confers a good outcome.

**Acknowledgments:** We thank Frank Doesburg for computer support and Iwan van der Horst for critically reading the manuscript.

**Author contributions:** EA conceived, helped execute the study and drafted the manuscript. BM helped execute the

study and drafted the manuscript. JD helped execute the study and drafted the manuscript. RP helped execute the study and revised the manuscript. AOL helped execute the study and drafted the manuscript. TL conceived, helped execute the study and drafted the manuscript. MN conceived, helped execute and drafted the manuscript. All authors read and approved final manuscript. All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

**Research funding:** None declared.

**Employment or leadership:** None declared.

**Honorarium:** None declared.

**Competing interests:** The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

## References

- Pickering BW, Gajic O, Ahmed A, Herasevich V, Keegan MT. Data utilization for medical decision making at the time of patient admission to ICU. *Crit Care Med* 2013;41:1502–10.
- Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute physiology and chronic health evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit Care Med* 2006;34:1297–310.
- Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, et al. SAPS 3 – from evaluation of the patient to evaluation of the intensive care unit. Part 2: development of a prognostic model for hospital mortality at ICU admission. *Intensive Care Med* 2005;31:1345–55.
- Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med* 1995;23:1638–52.
- Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, et al. The SOFA (sepsis-related organ failure assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med* 1996;22:707–10.
- Kouri T, Kairisto V, Virtanen A, Uusipaikka E, Rajamäki A, Finneman H, et al. Reference intervals developed from data for hospitalized patients: computerized method based on combination of laboratory and diagnostic data. *Clin Chem* 1994;40:2209–15.
- Finfer S, Wernerman J, Preiser JC, Cass T, Desai T, Hovorka R, et al. Clinical review: consensus recommendations on measurement of blood glucose and reporting glycemic control in critically ill adults. *Crit Care* 2013;17:229.
- Krinsley JS. Glycemic variability: a strong independent predictor of mortality in critically ill patients. *Crit Care Med* 2008;36:3008–13.
- Hessels L, Hoekstra M, Mijzen LJ, Vogelzang M, Dieperink W, Oude Lansink A, et al. The relationship between serum potassium, potassium variability and in-hospital mortality in critically ill patients and a before-after analysis on the impact of computer-assisted potassium control. *Crit Care* 2015;19:4.
- Oude Lansink-Hartgring A, Hessels L, Weigel J, de Smet AM, Gommers D, Nannan Panday PV, et al. Long term changes in dysnatremia incidence in the ICU: a shift from hyponatremia to hypernatremia. *Ann Intensive Care* 2016;6:22.
- Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *Int J Epidemiol* 1999;28:964–74.
- Hansen JB, Wilsgard L, Osterud B. Biphasic changes in leukocytes induced by strenuous exercise. *Eur J Appl Physiol Occup Physiol* 1991;62:157–61.
- Kilickaya O, Schmickl C, Ahmed A, Pulido J, Onigkeit J, Kashani K, et al. Customized reference ranges for laboratory values decrease false positive alerts in intensive care unit patients. *PLoS One* 2014;9:e107930.
- Stachon A, Segbers E, Hering S, Kempf R, Holland-Letz T, Krieg M. A laboratory-based risk score for medical intensive care patients. *Clin Chem Lab Med* 2008;46:855–62.
- Solinger AB, Rothman SI. Risks of mortality associated with common laboratory tests: a novel, simple and meaningful way to set decision limits from data available in the Electronic Medical Record. *Clin Chem Lab Med* 2013;51:1803–13.
- Bakker J, Nijsten MW, Jansen TC. Clinical use of lactate monitoring in critically ill patients. *Ann Intensive Care* 2013;3:12.
- Vincent JL, Quintairos E Silva A, Couto Jr. L, Taccone FS. The value of blood lactate kinetics in critically ill patients: a systematic review. *Crit Care* 2016;20:257.
- Jansen TC, van Bommel J, Schoonderbeek FJ, Sleswijk Visser SJ, van der Klooster JM, Lima AP, et al. Early lactate-guided therapy in intensive care unit patients: a multicenter, open-label, randomized controlled trial. *Am J Respir Crit Care Med* 2010;82:752–61.
- Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). *J Am Med Assoc* 2016;315:801–10.
- Addison DJ. Is routine ordering of both hemoglobin and hematocrit justifiable? *Canad Med Ass J* 1966;95:974–5.
- Nijboer JM, van der Horst IC, Hendriks HG, ten Duis HJ, Nijsten MW. Myth or reality: hematocrit and hemoglobin differ in trauma. *J Trauma* 2007;62:1310–2.
- Sakr Y, Rother S, Ferreira AM, Ewald C, Dünisch P, Riedemann N, et al. Fluctuations in serum sodium level are associated with an increased risk of death in surgical ICU Patients. *Critical Care Med* 2013;41:133–42.
- Whitfield JB. Gamma glutamyltransferase. *Crit Rev Clin Lab Sci* 2001;38:263–355.
- Stenius U, Rubin K, Gullberg D, Hogberg J. Gamma-glutamyl-transpeptidase-positive rat hepatocytes are protected from GSH depletion, oxidative stress and reversible alterations of collagen receptors. *Carcinogenesis* 1990;11:69–73.
- Kobayashi H, Nonami T, Kurokawa T, Kitahara S, Harada A, Nakao A, et al. Changes in the glutathione redox system during ischemia and reperfusion in rat liver. *Scand J Gastroenterol* 1992;27:711–16.
- Rajpert-De Meyts E, Shi M, Robison TW, Groffen J, Heisterkamp N, Forman HJ. Transfection with gamma-glutamyl transpeptidase enhances recovery from glutathione depletion using extracellular glutathione. *Toxicol Appl Pharmacol* 1992;114:56–62.

27. Lee DS, Evans JC, Robins SJ, Wilson PW, Albano I, Fox CS, et al. Gamma glutamyltransferase and metabolic syndrome, cardiovascular disease, and mortality risk: the Framingham heart study. *Arterioscler Thromb Vasc Biol* 2007;27:127–33.
28. Alkozai EM, Lisman T, Porte RJ, Nijsten MW. Early elevated serum gamma glutamyl transpeptidase after liver transplantation is associated with better survival. *F1000Res* 2014;3:85.
29. Alkozai EM, Nijsten MW, de Jong KP, de Boer MT, Peeters PM, Slooff MJ, et al. Immediate postoperative low platelet count is associated with delayed liver function recovery after partial liver resection. *Ann Surg* 2010;251:300–6.
30. Haveman JW, Zeebregts CJ, Verhoeven EL, van den Berg P, van den Dungen JJ, Zwaveling JH, et al. Changes in laboratory values and their relationship with time after rupture of an abdominal aortic aneurysm. *Surg Today* 2008;38:1091–101.
31. Keays R, Harrison PM, Wendon JA, Forbes A, Gove C, Alexander GJ, et al. Intravenous acetylcysteine in paracetamol induced fulminant hepatic failure: a prospective controlled trial. *Br Med J* 1991;303:1026–9.
32. Heard KJ. Acetylcysteine for acetaminophen poisoning. *N Engl J Med* 2008;359:285–92.
33. Sklar GE, Subramaniam M. Acetylcysteine treatment for non-acetaminophen-induced acute liver failure. *Ann Pharmacother* 2004;38:498–500.
34. Van de Moortel L, Speeckaert MM, Fiers T, Oeyen S, Decruyenaere J, Delanghe J. Low serum creatine kinase activity is associated with worse outcome in critically ill patients. *J Crit Care* 2014;29:786–90.
35. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361–87.

---

**Supplemental Material:** The online version of this article offers supplementary material (<https://doi.org/10.1515/cclm-2016-1028>).